

Performance of ChatGPT on USMLE: Potential for AI-assisted medical education using large language models

Year: 2022 | Citations: 2906 | Authors: Tiffany H. Kung, Morgan Cheatham, Arielle Medenilla, Czarina Sillos, Lorie De Leon

Abstract

We evaluated the performance of a large language model called ChatGPT on the United States Medical Licensing Exam (USMLE), which consists of three exams: Step 1, Step 2CK, and Step 3. ChatGPT performed at or near the passing threshold for all three exams without any specialized training or reinforcement. Additionally, ChatGPT demonstrated a high level of concordance and insight in its explanations. These results suggest that large language models may have the potential to assist with medical education, and potentially, even clinical decision-making.